CLAIMS

1. A polypeptide comprising an enterokinase recognition sequence and having the formula:

(1) Z_1-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Asp-Arg-Xaa_5-Z_2 (SEQ ID NO:1),

wherein Xaa_1 is an optional amino acid residue which, if present, is Ala, Asp, Glu, Phe, Gly, Ile, Asn, Ser, or Val; Xaa_2 is an optional amino acid residue which, if present, is Ala, Asp, Glu, His, Ile, Leu, Met, Gln, or Ser; Xaa_3 is an optional amino acid residue which, if present, is Asp, Glu, Phe, His, Ile, Met, Asn, Pro, Val, or Trp; Xaa_4 is Ala, Asp, Glu, or Thr; and Xaa_5 can be any amino acid residue; and wherein Z_1 and Z_2 are both optional and are, independently, polypeptides of one or more amino acids.

2. The polypeptide of Claim 1, wherein Xaa_1 is Asp, Xaa_2 is Ile, Xaa_3 is Asn, Xaa_4 is Asp, and Xaa_5 is Met, Thr, Ser, Ala, Asp, Leu, Phe, Asn, Trp, Ile, Gln, Glu, His, Val, Gly, or Tyr.

3. The polypeptide of Claim 1, wherein Z_1 is a ligand recognition sequence.

4. The polypeptide of Claim 1, wherein Z_1 is a streptavidin binding domain.

5. The polypeptide of Claim 4, wherein the streptavidin binding domain is selected from the sequences: His-Pro-Gln-Phe (SEQ ID NO:6), Cys-His-Pro-Gln-Phe-Cys (SEQ ID NO:5), Cys-His-Pro-Gln-Phe-Cys-Ser-Trp-Arg (SEQ ID NO:7), Trp-His-Pro-Gln-Phe-Ser-Ser (SEQ ID NO:210), Pro-Cys-His-Pro-Gln-Phe-Pro-Arg-Cys-Tyr (SEQ ID NO:211), and tandemly arranged combinations and repeats thereof.

6. The polypeptide of Claim 1, wherein Z_2 is a protein of interest.

7. The polypeptide of Claim 1, wherein the polypeptide Xaa_5-Z_2 is a protein of interest.

8. A polypeptide comprising an enterokinase recognition sequence and having the formula:

(2) Z_1-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Glu-Arg-Xaa_5-Z_2 (SEQ ID NO:2),

wherein Xaa_1 is an optional amino acid residue which, if present, is Asp or Glu; Xaa_2 is an optional amino acid residue which, if present, is Val; Xaa_3 is an optional amino acid residue which, if present, is Tyr; Xaa_4 is Asp, Glu, or Ser; and Xaa_5 can be any amino

acid residue; and wherein Z_1 and Z_2 are both optional and are, independently, polypeptides of one or more amino acids.

9. The polypeptide of Claim 8, wherein Xaa_5 is Met, Thr, Ser, Ala, Asp, Leu, Phe, Asn, Trp, Ile, Gln, Glu, His, Val, Gly, or Tyr.

10. The polypeptide of Claim 8, wherein Z_1 is a ligand recognition sequence.

11. The polypeptide of Claim 8, wherein Z_1 is a streptavidin binding domain.

12. The polypeptide of Claim 9, wherein the streptavidin binding domain is selected from the sequences: His-Pro-Gln-Phe (SEQ ID NO:6), Cys-His-Pro-Gln-Phe-Cys (SEQ ID NO:5), Cys-His-Pro-Gln-Phe-Cys-Ser-Trp-Arg (SEQ ID NO:7), Trp-His-Pro-Gln-Phe-Ser-Ser (SEQ ID NO:210), Pro-Cys-His-Pro-Gln-Phe-Pro-Arg-Cys-Tyr (SEQ ID NO:211), and tandemly arranged combinations and repeats thereof.

13. A polypeptide comprising an enterokinase recognition sequence having a sequence selected from the group consisting of SEQ ID NOs: 10 - 73 and 75 - 193, as shown in Tables 1, 2, 3, and 4.

14. A polynucleotide encoding an enterokinase cleavable fusion protein including the following domains, arranged in the direction of amino-terminus to carboxy-terminus: a ligand recognition sequence, an enterokinase recognition sequence having the formula Asp-Ile-Asn-Asp-Asp-Arg (SEQ ID NO:208) or Gly-Asn-Tyr-Thr-Asp-Arg (SEQ ID NO:209), and a protein of interest.

15. A vector comprising circular DNA and including the polynucleotide of Claim 14.

16. An expression vector comprising the polynucleotide of Claim 14 operably linked to a promoter sequence for expression in a recombinant host cell.

17. The expression vector of Claim 16, further comprising a signal sequence operably linked to the polynucleotide for effecting secretion of the expressed fusion protein into a culture medium.

18. A host cell transformed with the vector according to Claim 16 or 17.

19. The host cell of Claim 18, wherein the host cell is prokaryotic.

20. The host cell of Claim 18, wherein the cell eukaryotic.

21. A method for isolating a protein of interest comprising:

(a) culturing a recombinant host cell expressing a recombinant polynucleotide encoding an enterokinase cleavable fusion protein including the following domains, arranged in the direction of amino-terminus to carboxy-terminus: a ligand recognition sequence, an enterokinase recognition sequence having the formula:

(1) Xaa_1-Xaa_2-Xaa_3-Xaa_4-Asp-Arg-Xaa_5 (SEQ ID NO:206),

wherein Xaa_1 is an optional amino acid residue which, if present, is Ala, Asp, Glu, Phe, Gly, Ile, Asn, Ser, or Val; Xaa_2 is an optional amino acid residue which, if present, is Ala, Asp, Glu, His, Ile, Leu, Met, Gln, or Ser; Xaa_3 is an optional amino acid residue which, if present, is Asp, Glu, Phe, His, Ile, Met, Asn, Pro, Val, or Trp; Xaa_4 is Ala, Asp, Glu, or Thr; and Xaa_5 can be any amino acid residue; or

(2) Xaa_1-Xaa_2-Xaa_3-Xaa_4-Glu-Arg-Xaa_5 (SEQ ID NO:207),

wherein Xaa_1 is an optional amino acid residue which, if present, is Asp or Glu; Xaa_2 is an optional amino acid residue which, if present, is Val; Xaa_3 is an optional amino acid residue which, if present, is Tyr; Xaa_4 is Asp, Glu, or Ser; and Xaa_5 can be any amino acid residue,

and a protein of interest, under conditions suitable for expression of said fusion protein;

(b) contacting the expressed fusion protein with a binding ligand immobilized on a solid support under conditions suitable for formation of a binding complex between the binding ligand and the ligand recognition sequence;

(c) contacting the binding complex with enterokinase; and

(d) recovering the protein of interest.

22. The method of Claim 21, further comprising:
step (a1) after step (a), wherein said fusion protein is not secreted on expression, of lysing the host cells, and (a2) separating the cellular debris from the culture medium.

23. The method of Claim 21, further comprising:
step (a1) after step (a), wherein said fusion protein is secreted on expression, of collecting culture media containing the secreted fusion protein.

24. The method according to Claim 21, wherein said fusion protein has the formula:

(1) Z_1-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Asp-Arg-Xaa_5-Z_2 (SEQ ID NO:1),

wherein Xaa_1 is an optional amino acid residue which, if present, is Ala, Asp, Glu, Phe, Gly, Ile, Asn, Ser, or Val; Xaa_2 is an optional amino acid residue which, if present, is Ala, Asp, Glu, His, Ile, Leu, Met, Gln, or Ser; Xaa_3 is an optional amino acid residue which, if present, is Asp, Glu, Phe, His, Ile, Met, Asn, Pro, Val, or Trp; Xaa_4 is Ala, Asp, Glu, or Thr; and Xaa_5 can be any amino acid residue; Z_1 is a polypeptide comprising the sequence His-Pro-Gln-Phe-Ser-Ser-Pro-Ser-Ala-Ser-Arg-Pro-Ser-Glu-Gly-Pro-Cys-His-Pro-Gln-Phe-Pro-Arg-Cys-Tyr-Ile-Glu-Asn-Leu-Asp-Glu-Phe-Ser-Gly-Leu-Thr-Asn-Ile (SEQ ID NO:84), and Xaa_5-Z_2 is a protein of interest.

25. The method according to Claim 21, wherein said fusion protein has the formula:

(2) Z_1-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Glu-Arg-Xaa_5-Z_2 (SEQ ID NO:2),

wherein Xaa_1 is an optional amino acid residue which, if present, is Asp or Glu; Xaa_2 is an optional amino acid residue which, if present, is Val; Xaa_3 is an optional amino acid residue which, if present, is Tyr; Xaa_4 is Asp, Glu, or Ser; and Xaa_5 can be any amino acid residue; Z_1 is a polypeptide comprising the sequence His-Pro-Gln-Phe-Ser-Ser-Pro-Ser-Ala-Ser-Arg-Pro-Ser-Glu-Gly-Pro-Cys-His-Pro-Gln-Phe-Pro-Arg-Cys-Tyr-Ile-Glu-Asn-Leu-Asp-Glu-Phe-Ser-Gly-Leu-Thr-Asn-Ile (SEQ ID NO:84), and Xaa-Z_5 is a protein of interest.

26. The method according to Claim 24, wherein Xaa_5 is Met, Thr, Ser, Ala, Asp, Leu, Phe, Asn, Trp, Ile, Gln, Glu, His, Val, Gly, or Tyr.

27. A method for isolating a genetic package of interest comprising the steps:

(a) expressing in a genetic package a fusion protein comprising a protein of interest fused to an enterokinase cleavage sequence fused to a polypeptide expressed on the surface of said genetic package;

(b) contacting the genetic package with a ligand for the protein of interest, which ligand is capable of being immobilized on a solid support, under conditions suitable for the formation of a binding complex between said ligand and said protein of interest;

(c) immobilizing said ligand on a solid support, either before or after said contacting step (b),

(d) contacting the immobilized binding complex formed in step (b) with enterokinase; and

(e) recovering the genetic package of interest from said solid support.

28. The method of Claim 27, wherein the ligand is biotinylated and the immobilization is by binding to immobilized streptavidin or avidin.

29. The method of Claim 27, wherein the ligand is immobilized by binding to an immobilized antibody that binds said ligand.

30. The method according to Claim 27, further comprising the step, after step (b):

(b1) washing the support to remove unbound materials.

31. The method according to Claim 27, wherein said protein of interest is an antibody or fragment thereof.

32. The method according to Claim 27, wherein said recovered genetic package is amplified in a host selected from the group consisting of bacterial cells, insect cells, mammalian cells, and yeast.

33. The method according to Claim 27, wherein said genetic package is selected from the group consisting of: bacteriophage, bacteria, bacterial spores, yeast cells, yeast spores, insect cells, eukaryotic viruses, and mammalian cells.

34. The method according to Claim 33, wherein said genetic package is a filamentous bacteriophage and the polypeptide expressed on the surface of said host is selected from the group consisting of: gene III protein (SEQ ID NO:213); domain 2::domain 3::transmembrane domain::intracellular domain of gene III protein (SEQ ID NOs:215); and domain 3::transmembrane domain::intracellular anchor of gene III protein (SEQ ID NOs:217).

35. The method according to Claim 33, wherein said genetic package is an M13 phage.

36. A method for controlling the activity of a protein of interest comprising the steps:
(a) expressing in a recombinant host a fusion protein comprising the elements (i) a first protein fused to (ii) an enterokinase cleavage sequence fused to (iii) a second protein, wherein said fusion protein has suppressed activity due to the conformation of elements (i), (ii) and (iii);
(b) treating the fusion protein with enterokinase such that said first protein and second protein are separated and at least one of said first protein and said second protein thereby exhibits the activity of a protein of interest.

37. The method according to Claim 36, wherein said second protein is the protein of interest and is a protease, and wherein said first protein is an inhibitor of the protease.

38. The method according to Claim 36, wherein said first protein is the protein of interest and is a protease, and wherein said second protein is an inhibitor of the protease.

39. The method according to Claim 36, wherein said first protein is the variable light (V_L) domain of an scFv antibody, and said second protein is the variable heavy (V_H) domain of an scFv antibody, and wherein said protein of interest is the scFv formed by the association of said first protein with said second protein.

40. The method according to Claim 36, wherein said second protein is the variable light (V_L) domain of an scFv antibody, and said first protein is the variable heavy (V_H) domain of an scFv antibody, and wherein said protein of interest is the scFv formed by the association of said first protein with said second protein.

41. A method for detecting the expression of a fusion protein on the surface of a recombinant host comprising the steps:

(a) expressing, in a recombinant host, a fusion protein comprising a first protein fused to an enterokinase cleavage sequence fused to a second protein fused to a polypeptide expressed on the surface of said host;

(b) contacting the host with a ligand for said first protein immobilized on a solid support under conditions suitable for forming a binding complex between the ligand and the first protein;

(c) removing unbound materials;

(d) treating any bound complex with enterokinase;

(e) recovering hosts released from said solid support, wherein said recovered hosts are verified expressors of said fusion protein.

42. The method according to Claim 41, wherein said second protein is an antibody or antibody fragment.

43. The method according to Claim 41, wherein said first protein is a streptavidin-binding polypeptide and said ligand is streptavidin.

44. A method of selecting display polypeptides from a display library that have specific affinity for a target, comprising the steps:

(a) providing a display library of polypeptides comprising a multiplicity of genetic packages, wherein each genetic package expresses a fusion protein that comprises an enterokinase recognition sequence between a diplay polypeptide library member and a polypeptide that anchors the fusion protein to the genetic package,

(b) contacting the display library with a target,

(c) immobilizing the target on a solid support, either before or after said contacting step (b),

(d) separating non-binding genetic packages from bound genetic packages,

(e) treating the bound genetic packages with enterokinase, and

(f) recovering and amplifying the genetic packages released.

45. The method of Claim 44, wherein the genetic package is an M13 phage.

46. The method of Claim 45, wherein the polypeptide that anchors the fusion protein to the genetic package comprises at least the domain 3::transmembrane domain::intracellular domain portion of the gene III protein.

47. The method of Claim 44, wherein the display polypeptides comprise human Fabs.

48. The method of Claim 44, wherein the display polypeptides comprise peptides of ten to twenty-one amino acids in length.

49. The method of Claim 48, wherein each peptide contains two cysteines.